Journal of Multivariate Analysis 154 (2017) 135–146

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Functional Cramér–Rao bounds and Stein estimators in Sobolev spaces, for Brownian motion and Cox processes

Eni Musta [a], Maurizio Pratelli [b], Dario Trevisan [b,*]

[a] Delft University of Technology, Netherlands

[b] Università degli Studi di Pisa, Italy

ARTICLE INFO

Article history:
Received 20 July 2015
Available online 29 October 2016

AMS 2010 subject classifications:
primary 62J07
secondary 60J65

Keywords:
Cramer–Rao bound
Stein phenomenon
Malliavin calculus
Cox model

ABSTRACT

We investigate the problems of drift estimation for a shifted Brownian motion and intensity estimation for a Cox process on a finite interval $[0, T]$, when the risk is given by the energy functional associated to some fractional Sobolev space $H_0^1 \subset W^{\alpha,2} \subset L^2$. In both situations, Cramér–Rao lower bounds are obtained, entailing in particular that no unbiased estimators (not necessarily adapted) with finite risk in H_0^1 exist. By Malliavin calculus techniques, we also study super-efficient Stein type estimators (in the Gaussian case).

1. Introduction

In this paper, we focus on two problems of nonparametric (or, more rigorously, infinite-dimensional parametric) statistical estimation: drift estimation for a shifted Brownian motion and intensity estimation for a Cox process, on a finite time interval $[0, T]$. Our investigation stems from the articles [10,11], where N. Privault and A. Réveillac developed an original approach to these problems, by employing techniques from Malliavin calculus to study Cramér–Rao bounds and super-efficient "shrinkage" estimators, originally developed by C. Stein in [5] and then expanded in [13], to fit in infinite-dimensional frameworks. Such a combination of these two powerful techniques can be cast into a more general picture, where Malliavin calculus tools provide insights in statistics and more generally, on probabilistic approximations: let us mention here the monograph [8], which collects many results of the fruitful meeting of another great contribution of C. Stein (the so-called Stein method) with Malliavin calculus, and other recent articles such as [2,4,7,12].

As in [10,11], here we assume that the unknown function to be estimated belongs to the Hilbert space $H_0^1(0, T)$ (which is a reasonable choice, at least in the case of shifted Brownian motion, because of the Cameron–Martin and Girsanov theorems) but we move further by addressing the following question, which is rather natural but has apparently not yet been considered: what about estimators that also take values in H_0^1? Indeed, in [10,11], estimators are seen as functions with values in $L^2([0, T], \mu)$ (where μ is any finite measure) or, equivalently, the associated risk is computed with respect to the L^2 norm and not the (stronger) H_0^1 norm.

To investigate this problem, we first provide Cramér–Rao bounds with respect to different risks, by considering the estimation in the interpolating fractional Sobolev space $H_0^1 \subset W^{\alpha,2} \subset L^2$, for $\alpha \in [0, 1]$. It turns out that no unbiased

* Corresponding author.
E-mail addresses: e.musta@tudelft.nl (E. Musta), pratelli@dm.unipi.it (M. Pratelli), dario.trevisan@unipi.it (D. Trevisan).

http://dx.doi.org/10.1016/j.jmva.2016.10.011
0047-259X/

estimator exists in H_0^1 (Theorem 2.5) and even in $W^{\alpha,2}$, for $\alpha \geq 1/2$ (Theorem 2.9). Although a bit surprising, these results reconcile with the following intuition: since the estimator is a function of the realization of the process, whose paths also do not belong to H_0^1 (nor $W^{\alpha,2}$, for $\alpha \geq 1/2$), it is "too risky" to estimate (without bias) the parameter on that scale of regularity. Therefore, besides answering a rather natural question, our results highlight the delicate role played by the choice of different norms in such estimation problems, and one might expect that similar phenomena might appear in other situations, technically more demanding, e.g., stochastic differential equations.

As a second task, we study super-efficient "shrinkage" estimators in the spaces $W^{\alpha,2}$. It is often suggested on heuristic grounds that the ideal situation for the problem of estimation would be to have an unbiased estimator with low variance, but that allowing for a little bias may allow one to find estimators with lower risks, in many situations: we strongly rely on the recent extensions and combinations of the original approach by Stein with Malliavin calculus to these frameworks developed in [10,11]. Using a similar approach, we give sufficient conditions for the existence of super-efficient estimators in $W^{\alpha,2}$, for $\alpha < 1/2$, and we give explicit examples of such estimators, in the case of Brownian motion (Example 3.3). In the case of Cox processes, although it is possible to define a suitable version of Malliavin calculus and provide sufficient conditions for Stein estimators, we are currently unable to provide explicit examples.

The paper is organized as follows. In Section 2 we deal with drift estimation for a shifted Brownian motion, addressing Cramér–Rao lower bounds with respect to risks computed in H_0^1 and fractional Sobolev spaces. In Section 3, we discuss super-efficient estimators. Finally, analogous results on intensity estimators for Cox processes are given in Section 4.

2. Drift estimation for a shifted Brownian motion

In this section, we fix $T \geq 0$ and let $X = (X_t)_{t\in[0,T]}$ be a Brownian motion (on the finite interval $[0, T]$), defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$. Instead of choosing a fixed (infinite-dimensional) space of parameters Θ, we simply notice that our arguments apply to any set Θ of absolutely continuous, adapted processes $u_t = \int_0^t \dot{u}_s ds$ (for $t \in [0, T]$) such that

(1) $(\dot{u}_t)_{t\in[0,T]}$ satisfies the conditions of Girsanov's theorem;
(2) Θ contains the Cameron–Martin space H_0^1;
(3) for any $u \in \Theta$, $v \in H_0^1$, one has $u + v \in \Theta$.

Let us recall that H_0^1 $(= H_0^1(0, T))$ is defined as the space of (continuous) functions of the form $h(t) = \int_0^t \dot{h}(s)ds$, for $t \in [0, T]$, with $\dot{h} \in L^2(0, T)$. In particular, we may let $\Theta = H_0^1$.

For $u \in \Theta$, we define the probability measure $P^u = L^u P$, with

$$L^u = \exp\left(\int_0^T \dot{u}_s dX_s - \frac{1}{2}\int_0^T \dot{u}_s^2 ds\right).$$

Girsanov's theorem entails that, with respect to the probability measure P^u, the process $X_t^u = X_t - u_t$ is a Brownian motion on $[0, T]$.

We address the problem of estimating the drift with respect to P^u on the basis of a single observation of X (of course, repeated and independent observations can improve the estimates, but this amounts to a simple generalization). Such a problem is of interest in different fields of application: for example, we can interpret X as the observed output signal of some unknown input signal u, perturbed by a Brownian noise. Such a problem is investigated, e.g., in [10], where the following definition is given.

Definition 2.1. Any measurable stochastic process $\xi : \Omega \times [0, T] \to \mathbb{R}$ is called an estimator of the drift u. An estimator of the drift u is said to be unbiased if, for every $u \in \Theta$, $t \in [0, T]$, ξ_t is P^u-integrable and one has $\mathrm{E}^u(\xi_t) = \mathrm{E}^u(u_t)$.

In this section, we forego the specification of "the drift u" and simply refer to estimators. Moreover, we refer to the quantity $\mathrm{E}^u(\xi_t - u_t)$ as the bias of the estimator ξ (whenever it is well-defined).

By introducing as a risk associated to any estimator ξ the quantity

$$\mathrm{E}^u(\|\xi - u\|^2_{L^2(\mu)}) = \mathrm{E}^u\left\{\int_0^T |\xi_t - u_t|^2 \mu(dt)\right\}, \tag{1}$$

where μ is any finite Borel measure on $[0, T]$, Privault and Réveillac provide the Cramér–Rao lower bound stated next for adapted and unbiased estimators [10, Proposition 2.1]. In what follows, Θ being the space of all absolutely continuous, adapted processes, whose derivatives satisfy the conditions of Girsanov's theorem.

Theorem 2.2 (*Cramér–Rao Inequality in $L^2(\mu)$*). *For any adapted and unbiased estimator ξ, one has*

$$\mathrm{E}^u(\|\xi - u\|^2_{L^2(\mu)}) \geq \int_0^T t\mu(dt), \quad \text{for every } u \in \Theta. \tag{2}$$

Equality is attained by the (efficient) estimator $\hat{u} = X$.